

The consumption-based determinants of the term structure of discount rates¹

Christian Gollier
University of Toulouse and EIF

April 30, 2007

¹An earlier version of this paper was entitled "Transitory shocks to GNP and the consumption-based term structure of interest rates". I am indebted to John Campbell, Martin Weitzman and to two referees for helpful comments.

Abstract

The rate of return of a zero-coupon bond with maturity T is determined by our expectations about the mean (+), variance (-) and skewness (+) of the growth of aggregate consumption between 0 and T . The shape of the yield curve is thus determined by how these moments vary with T . We first examine growth processes in which a higher past economic growth yields a first-degree dominant shift in the distribution of the future economic growth, as assumed for example by Vasicek (1977). We show that when the growth process exhibits such a positive serial dependence, then the yield curve is decreasing if the representative agent is prudent ($u''' > 0$), because of the increased risk that it yields for the distant future. A similar definition is proposed for the concept of second-degree stochastic dependence, as observed for example in the Cox-Ingersoll-Ross model, with the opposite comparative static property holding under temperance ($u'''' < 0$), because the change in downside risk (or skewness) that it generates. Finally, using these theoretical results, we propose two arguments in favor of using a smaller rate to discount cash-flows with very large maturities, as those associated to global warming or nuclear waste management.

Keywords: Stochastic dependence, yield curve, far distant future, prudence, temperance, downside risk.

JEL Classification: G12, E43, Q51

1 Introduction

How much effort are we ready to make today to improve the future? Households are faced with this question when they plan their savings for retirement, whereas entrepreneurs have to determine whether to undertake new investment projects. At the collective level, one needs to determine, for example, whether to limit the national budget deficit, or whether to invest in the education system. In a recent past, similar questions emerged, but with the striking innovation of being related to the far-distant future. Exploring the universe, protecting the biodiversity, limiting the extraction of exhaustible resources, dealing with nuclear wastes and global warming are a few examples of policy questions that confront us to our attitude towards improving the welfare of human beings that will live in hundreds or thousands years in the future. These valuation questions are all solved by the selection of the discount rate.

As is well-known, the use of a single rate to discount sure cash-flows at all maturities implies that costs and benefits occurring, say, in more than 100 years are typically irrelevant for the decision, because of the exponential nature of discounting. This is why for example the so-called "Copenhagen Consensus"¹ ranked all projects linked to the prevention of global warming at the lowest priority level based on standard cost-benefit analyses with a constant discount rate. The problem is that there is a priori no scientific reason to believe that one should discount all maturities at the same rate. The tradition of using a constant rate in cost-benefit analysis should not be seen as a dogma, but rather as a useful practical simplification. Various authors – among whom Weitzman (1998, 2001, 2004) is the most vocal – claimed that one should opt for discount rates that are decreasing with the maturity of the cash flows under scrutiny. Weitzman (2004) in particular develops an argument for selecting a zero discount rate for maturities around 50 years, the discount rate becoming even negative for longer time horizons. Of course, adopting such recommendations would massively reallocate our collective investments towards those benefiting to distant generations, potentially at the detriment of actions with more immediate benefits such as fighting malaria and promoting education in developing countries. It is therefore important

¹It is the outcome of a conference held in Copenhagen in May 2004 aimed at ranking a set of various collective investment projects, including fighting AIDS and malaria in developing countries, water management, biodiversity, education,....

to have a good understanding of the reasons why we should adopt such decreasing discount rates.

Since the seminal contribution of Vasicek (1977), economists have intensively explored how efficient discount rates should vary with the maturity of the corresponding cash payment. The immense literature on the term structure of interest rates has produced an important corpus of knowledge about this question. It is quite unfortunate that researchers discussing this question in the various forums of environmental economics do not take advantage of the existence of this vast literature.² There are several reasons for that. First, most papers on the yield curve are aimed at explaining the observed shape of that curve, whereas environmental economists have a much more normative approach. Notice however that the absence of frictions in the standard models on the term structure implies that the equilibrium interest rates are also the socially efficient discount rates to be used in cost-benefit analysis. Second, researchers in finance are usually interested in pricing traded assets, which implies that their time-horizon is limited by the largest maturity of existing liquid markets for risk-free assets, which does not exceed 30 years. Last but not least, this literature is highly complex, and it does usually not provide intuition to the underlying phenomena. This is well summarized by Piazzesi (2005): "The quest for understanding what moves bond yields has produced an enormous literature with its own journals and graduate courses. Those who want to join the quest are faced with considerable obstacles. The literature has evolved mostly in continuous time, where stochastic calculus reigns and partial differential equations spit fire. The knights in this literature are fighting for different goals, which makes it often difficult to comprehend why the quest is moving in certain directions." This quest leads to the (preliminary) conclusion that the shape of the yield curve is governed by the dynamics of the short term interest rate (and maybe a few other stochastic factors) that may entails mean reversion together with temporary and permanent shocks. Because the term structure is obtained by arbitrage using an exogenously given dynamic process for the price kernel, this reasoning is usually not based on individual preferences. It is therefore not an easy starting block to explain to public decision-makers how much effort our

²See for example the collective book edited by Portney and Weyant (1999) on discounting. See also Arrow et al. (1996), Weitzman (1998, 2001), Newell and Pizer (2003) and Groom, Koundouri, Panopoulou and Pantelidis (2004).

generation should undertake to improve the welfare of future generations.

The aim of this paper is twofold. First, we exhibit the fundamental determinants of the shape of the yield curve based on the preferences of the representative agent and on the stochastic process of aggregate consumption in the economy. Second, we examine realistic dynamic growth processes that are relevant to determine the very long discount rates. We consider the classical Lucas (1978)'s tree economy with an exogenous growth process to examine these questions.

The efficient interest rate associated to time horizon t is decreasing in our willingness to save in order to finance consumption at that date, which itself depends upon our expectations about the growth of our incomes over $[0, t]$. Therefore, the term structure of interest rates provides a rich set of information about these expectations. For example, when consumers expect an increase in their future incomes, they want to cash this benefit immediately by reducing their saving. This raises the equilibrium interest rate. This wealth effect relies on the standard assumption that consumers want to smooth their consumption over time. It explains why the yield curve is upward sloping when the representative agent expects an accelerating growth rate in the future (Estrella and Hardouvelis (1991)).

Among the many difficulties to extract testable hypothesis about the relationship between the term structure and expectations about the future economic activity, the most important one is due to uncertainty. Since Leland (1968), we know that uncertainty about future incomes raises the prudent consumers' willingness to save. This precautionary effect tends to reduce the interest rate. This implies for example that the anticipation of a deterministic reduction in the volatility of growth yields an increasing yield curve (Barsky (1989)). It is interesting to examine how does the accumulation of risk for longer time horizons influence the determination of the corresponding interest rate. Because longer horizons mean larger expected consumption, people want to save less for these better times. On the contrary, longer horizons also mean more risk, which implies that consumers want to save more for these more uncertain times. Which of these wealth and precautionary effects will dominate the other? If the wealth effect dominates the precautionary effect, then the yield curve must be increasing.

The simplest case is when the growth of the economy follows a stationary random process. In this case, both the expected log consumption and its variance increases proportionally with the time horizon. It implies that

the wealth effect and the precautionary effect exactly compensate each other when the representative agent has a constant relative risk aversion (CRRA). As is well-known (see for example Mankiw (1981)), CRRA combined with an i.i.d. consumption growth process implies that the yield curve is completely flat. In sections 3 and 4 of this paper, we show how the existence of serial correlations in the growth rate of the economy affects the shape of the yield curve. We define two types of serial correlations. Positive first-degree stochastic dependence (FSD) occurs when an increase in the first subperiod growth rate induces a first-degree stochastic improvement in the conditional distribution of the growth rate in the second subperiod. Such a positive serial dependence in the growth of the economy tends to magnify the long-term risk on consumption relative to the short-term risk. It implies that the prudent representative agent will want to rebalance her efforts towards the longer time horizons, thereby tending to reduce long interest rates. This is formally shown in section 3 in a much simpler and more intuitive way than traditionally done in the existing literature. It is also more general in the sense that our result only requires that the representative agent be prudent. FSD dependence is the main feature of the two classical models of the term structure, namely Vasicek (1977) and Cox, Ingersoll and Ross (1985a,b).

There is positive second-degree stochastic dependence (SSD) in growth rates if an increase in the first subperiod rate yields an increase in risk in the conditional distribution of growth in the second subperiod. This tends to raise the skewness of the distribution of future consumption. Ex ante, it reduces the expected marginal utility of wealth at that maturity if the fourth derivative of the utility function is negative, a condition that is satisfied for CRRA preferences. This tends to reduce the willingness to purchase more zero-coupon bonds associated to long maturities, thereby raising their rate of return. This is proved in section 4. Notice that the main feature of the Cox-Ingersoll-Ross model is to add some SSD dependence in the Vasicek model. The link of our results to these two classical models are made more explicit in section 5.

In section 6, we examine two specific stochastic processes with positive FSD dependence that are realistic representations of the uncertainty faced by Humanity in the very long run. The first stochastic process for aggregate consumption has a drift that can take two possible values. A switch from one drift to the other can occur at each period with a very small probability. This is aimed at modeling the kind of event that we experienced with the

industrial revolution at the end of the 18th century, where the drift changed quite abruptly from the secular 0% per year to 2% per year since then. Our model formalizes the risk of a switch in the opposite direction – ”The Limit to Growth” – due for example to the scarcity of natural resources or to the extinction of scientific progresses. We show that the positive FSD dependence that this stochastic process yields a strong negative effect on the rate at which we should discount far-distant cash-flows. In the second model of long-term uncertainty inspired from Weitzman (2004), we assume that the drift is unique but unknown. As time goes by, one will use Bayes rule to update the beliefs about the true value of the drift. This stochastic process also yields positive FSD dependence – and thus decreasing discount rates – for the simple reason that a good news in the short term is a good news for the secular distribution of growth.

2 The term structure

The preferences of the representative agent in the economy are represented by her utility function u and by her rate of pure preference for the present δ . The utility function u on consumption is assumed to be three times differentiable, increasing and concave. Let \tilde{c}_t denote consumption at date t . The equilibrium per period rate of return at date 0 for a zero-coupon bond maturing at date t is denoted r_t . To be in equilibrium, investing marginally in such an asset should leave the expected discounted utility of the representative agent unchanged. This condition is written as

$$e^{-\delta t} E u'(\tilde{c}_t) e^{r_t t} = u'(c_0), \quad (1)$$

which is the standard Euler equation for the consumption-saving problem. On the right-hand side of this equality, $u'(c_0)$ is the welfare cost of reducing consumption by one monetary unit, which is invested in the zero-coupon bond. The left-hand side is the welfare benefit that such investment yields. Consumption at date t is increased by $e^{r_t t}$, which yields an increase in expected utility by $E u'(\tilde{c}_t) e^{r_t t}$, which must be discounted at rate δ to take account of the delay. The classical consumption-based pricing formula is

obtained by rewriting condition (1) as

$$r_t = \delta - \frac{1}{t} \ln \frac{Eu'(\tilde{c}_t)}{u'(c_0)}. \quad (2)$$

Two factors determine by how much the risk-free rate exceeds the rate of pure preference for the present δ . The first factor is a wealth effect. If we expect to consume more in the future, i.e., if $E\tilde{c}_t > c_0$, the marginal utility of one more euro in the future is smaller than the marginal utility of one more euro immediately: $u'(E\tilde{c}_t) < u'(c_0)$. It implies that $-t^{-1} \ln(u'(E\tilde{c}_t)/u'(c_0))$ is positive. This positive wealth effect is increasing in the expected growth rate of consumption over the entire period $[0, t]$ and in the rate at which marginal utility is decreasing with consumption, which is measured by the index of relative risk aversion $R(c) = -cu''(c)/u'(c)$. The intuition is that higher expectations about future incomes reduces the willingness to save, thereby raising the equilibrium interest rate.

But, except when marginal utility is linear, $Eu'(\tilde{c}_t)$ is not equal to $u'(E\tilde{c}_t)$, which introduces a second factor to the determination of interest rates. When the representative agent is prudent, i.e., when marginal utility is convex, the uncertainty surrounding future consumption raises the expected marginal utility: $Eu'(\tilde{c}_t) > u'(E\tilde{c}_t)$. This raises the willingness to save, thereby yielding a reduction of the equilibrium interest rate. This precautionary effect goes opposite to the wealth effect. It is increasing in the riskiness of future consumption and in the index of convexity of marginal utility, which is defined as relative prudence $P(c) = -cu'''(c)/u''(c)$. We can make these different factors more explicit by using second-order Taylor approximations of $u'(z_t)$ in the above equality. This technique yields

$$r_t \simeq \delta + R(c_0) \frac{E\tilde{c}_t - c_0}{tc_0} - \frac{1}{2} R(c_0) P(c_0) \frac{Var(\tilde{c}_t/c_0)}{t}, \quad (3)$$

where the three terms in the right-hand side measure respectively the impatience effect, the wealth effect and the precautionary effect. This approximation is exact for the instantaneous rate r_0 .

The term structure of interest rate is determined by how these two conflicting factors are compounded over time. A more distant future usually yields a larger expected consumption and a larger uncertainty. The risk-averse and prudent representative agent's willingness to purchase zero-

coupon bonds with that long maturity is reduced by the larger expected consumption, and is increased by the larger uncertainty. Therefore, as suggested by approximation (3), an increasing (decreasing) yield curve is obtained if the wealth effect becomes more (less) dominant compared to the precautionary effect when considering longer time horizons.

To illustrate, let us consider a simple case. Suppose that $u(c) = c^{1-\gamma}/(1-\gamma)$, which implies that $R(c) = \gamma$ and $P(c) = \gamma + 1$ for all c . Suppose also that the logarithm of consumption follows a stationary Brownian motion:³

$$d \ln c_t = \mu dt + \sigma dz_t, \quad (4)$$

where μ and σ are two scalars measuring respectively the mean and standard deviation of the change in log consumption.

Proposition 1 *Suppose that relative risk aversion is a constant γ and that the log of consumption follows a stationary Brownian motion with trend μ and volatility σ . Then, the yield curve is flat with $r_t = r_0 = \delta + \gamma\mu - 0.5\gamma^2\sigma^2$ for all t .*

Proof: Because $u'(c) = c^{-\gamma}$, we have that

$$\frac{Eu'(\tilde{c}_t)}{u'(c_0)} = E \exp[-\gamma(\ln \tilde{c}_t - \ln c_0)]$$

By assumption, $\ln \tilde{c}_t - \ln c_0$ is normally distributed with mean μt and variance $\sigma^2 t$. We can thus rewrite the above equation as

$$\frac{Eu'(\tilde{c}_t)}{u'(c_0)} = \frac{1}{\sigma\sqrt{2\pi t}} \int \exp(-\gamma z) \exp\left(-\frac{(z - \mu t)^2}{2\sigma^2 t}\right) dz.$$

This can be rewritten as

$$\frac{Eu'(\tilde{c}_t)}{u'(c_0)} = \exp\left(-\gamma\left(\mu t - \frac{\gamma\sigma^2 t}{2}\right)\right) \left[\frac{1}{\sigma\sqrt{2\pi t}} \int \exp\left(-\frac{(z - (\mu t - \gamma\sigma^2 t))^2}{2\sigma^2 t}\right) dz\right].$$

The bracketed term is the integral of the normal density function with mean $\mu t - \gamma\sigma^2 t$ and variance $\sigma^2 t$. This equals unity. Thus we obtain that

$$\frac{Eu'(\tilde{c}_t)}{u'(c_0)} = \exp[-\gamma(\mu t - 0.5\gamma\sigma^2 t)].$$

³Using Ito's Lemma, this is equivalent to assume that $dc/c = (\mu + 0.5\sigma^2)dt + \sigma dz$.

Thus, using equation (2) yields⁴

$$r_t = \delta + \gamma\mu - 0.5\gamma^2\sigma^2. \quad \blacksquare \quad (5)$$

This formula is equivalent to those obtained by Mankiw (1981), Hansen and Singleton (1983), Breeden (1986) and Campbell (1986). It shows that when relative risk aversion is constant (CRRA) and the growth rate of the economy follows a stationary Brownian motion, a longer time horizon yields an increase in the wealth effect and an increase in the precautionary effect that exactly compensate each other, yielding a flat yield curve.

Gollier (2002 a,b) characterized the conditions on preferences that imply a monotone yield curve under the assumption of a stationary Brownian motion. For example, he shows that increasing relative risk aversion implies an increasing yield curve if the probability of recession is small enough. In this paper, we follow a more standard strategy which consists in relaxing the assumption of a stationary Brownian motion. This is relaxed by assuming that the mean μ and/or the volatility σ of the consumption growth process are path-dependent, i.e., that the growth at time t depends upon the growth in the periods preceding t . In a word, we assume that future growth rates are predictable. The typical methodology in the literature on the term structure of interest rates is to assume the following time series model:

$$\begin{aligned} d \ln c_t &= \mu(s)dt + \sigma_c(s)dz_t, \\ ds &= g(s)dt + \sigma_s(s)dz_t. \end{aligned}$$

Both the mean and the volatility of the growth rate of the economy are affected by a state variable (also called a "factor") s that itself follows a potentially non-stationary Brownian motion. The special case of a deterministic process for the state variable ($\sigma_s = 0$) is easy to treat using the above integration method. For example, when $\sigma_c(s) \equiv \sigma$, we easily obtain in the CRRA case that

$$r_t = \delta + \gamma m(t) - 0.5\gamma^2\sigma^2, \quad (6)$$

⁴We can reconcile equations (3) and (5) by observing that the growth rate of expected consumption equals $\mu + 0.5\sigma^2$. This implies that equation (5) can be rewritten as

$$r_t = \delta + R \frac{dE c_t}{c_t} - 0.5RP\sigma^2.$$

This proves that approximation (3) is exact in this case.

where $m(t) = t^{-1} \int_0^t \mu(s(\tau)) d\tau$ is the mean change in log consumption in period $[0, t]$, and $s(\tau)$ is the solution of the differential equation $s' = g(s)$ with initial condition $s(0) = s_0$. Only the wealth effect is affected by the deterministic change in the expectation μ about the growth rate of the economy. These changes in expectation explain why the yield curve is usually not flat for short and medium time-horizons. For example, the expectation of an accelerating growth implies an increasing yield curve. Observe from (6) that the unpredictable shocks in changes in log consumption have no effect on the shape of the yield curve. It only shifts it downwards.

The complexity of the theory on the yield curve comes from the stochastic component of the motion of the state variable s ($\sigma_s \neq 0$). In this paper, we isolate two effects of these predictable changes in expectations. Suppose first that the volatility σ_c of the growth rate of the economy is constant. When σ_c and σ_s have the same sign, and when μ is increasing in s , the expected future growth rate of consumption is positively correlated with the short-term growth rate. More precisely, an increase in the stochastic component dz_t of the short-term growth yields a first-degree stochastic dominant shift in the future growth rate. In section 3, we examine the effect of this positive correlation on the shape of the yield curve. Alternatively, suppose that the expected growth rate μ of the economy is state-independent, and that the volatility of the growth rate is increasing in the state variable. Then, if σ_c and σ_s have the same sign, the volatility of the future growth rate of the economy is positively correlated with the short-term growth rate. More precisely, an increase in the growth rate dz_t in the short run yields a second-degree stochastic shift in the future growth rate. We examine the effect of this type of statistical relations on the shape of the yield curve in section 4.

3 First-degree stochastic dependence

In this section, we consider an arbitrary stochastic process for \tilde{c}_t . We examine the effect of positive serial dependence in changes in consumption on the interest rate associated to maturity T . To do so, let us split period $[0, T]$ into two subperiods $[0, t]$ and $[t, T]$. Consider a random vector $(\tilde{x}_1, \tilde{x}_2)$ where \tilde{x}_1 and \tilde{x}_2 denote the change in consumption respectively in period $[0, t]$ and $[t, T]$. It implies that consumption at date T equals $c_0 + \tilde{x}_1 + \tilde{x}_2$. Let F denote the distribution function of $(\tilde{x}_1, \tilde{x}_2)$, and let F_1 and F_2 denote, respectively,

the marginal distributions of \tilde{x}_1 and \tilde{x}_2 . Let also $F_{2|1}$ be the conditional distribution of \tilde{x}_2 : $F_{2|1}(x_1, x_2) = \Pr[\tilde{x}_2 < x_2 \mid \tilde{x}_1 = x_1]$. We suppose that this distribution function exists.

Definition 1 Consider a pair of random variables $(\tilde{x}_1, \tilde{x}_2)$. We say that there is positive FSD dependence between \tilde{x}_1 and \tilde{x}_2 if $F_{2|1}$ is nonincreasing in x_1 for all x_2 .

In other words, an increase in x_1 generates a first-order stochastic dominant shift in the conditional distribution of \tilde{x}_2 . In the statistical literature, this notion is referred to as the "stochastic increasing positive dependence", because \tilde{x}_2 is more likely to take on larger value when x_1 increases (see for example Joe (1997)). Milgrom (1981) uses this concept to define the notion of a good news. An example of stochastic process that satisfies the FSD property is the AR(1) process $\tilde{x}_2 = \phi\tilde{x}_1 + \tilde{\varepsilon}$ with a positive ϕ .

The long-term interest rate in such an economy equals

$$r_T = \delta - \frac{1}{T} \ln \frac{Eu'(c_0 + \tilde{x}_1 + \tilde{x}_2)}{u'(c_0)}. \quad (7)$$

We want to compare this rate to the one that would prevail in an economy with the same marginal distributions for \tilde{x}_1 and \tilde{x}_2 , but with no serial dependence between them. In the economy without any serial dependence, the long-term interest rate would equal

$$r_T^i = \delta - \frac{1}{T} \ln \frac{Eu'(c_0 + \tilde{x}_1 + \tilde{x}_2^i)}{u'(c_0)}, \quad (8)$$

where $(\tilde{x}_1, \tilde{x}_2^i)$ is a vector of independent random variables with distribution F_1 and F_2 , respectively. Interest rate r_T^i would be what one would obtain from the calibrated model by assuming independence and by using the observed variance of annual changes in consumption as the estimation of $Var(\Delta c)$. We want to determine the conditions under which r_T is smaller than r_T^i , when \tilde{x}_2 exhibits positive FSD dependence with respect to \tilde{x}_1 . There is a simple intuition for why this should be the case. The existence of a positive dependence in the changes in consumption tends to magnify the long-term risk compared to short-term risks. This induces the prudent representative agent to purchase more zero-coupon bonds with a long maturity, thereby

reducing the equilibrium long-term rate. Comparing (7) and (8) implies that r_T is smaller than r_T^i if

$$Eu'(c_0 + \tilde{x}_1 + \tilde{x}_2) \geq Eu'(c_0 + \tilde{x}_1 + \tilde{x}_2^i). \quad (9)$$

The following Lemma is useful to examine this problem.

Lemma 1 *Consider a differentiable bivariate function h . The following two conditions are equivalent:*

1. *For any pair of random variables $(\tilde{x}_1, \tilde{x}_2)$ that satisfies positive first-order dependence, we have that*

$$Eh(\tilde{x}_1, \tilde{x}_2) \geq Eh(\tilde{x}_1, \tilde{x}_2^i) \quad (10)$$

2. *h is supermodular, i.e., $\partial h / \partial x_2$ is increasing in x_1 .*

Proof: See the appendix.⁵

Tchen (1980) showed a closely related result: If condition (10) is satisfied for all supermodular functions, then $F(x_1, x_2) \leq F_1(x_1)F_2(x_2)$ for all (x_1, x_2) , i.e., $(\tilde{x}_1, \tilde{x}_2)$ are "positive quadrant dependent", a concept weaker than FSD dependence.⁶ Shaked and Shanthikumar (2007) generalize these properties of "supermodular orders" to more than two random variables.

Applying this lemma to condition (9) requires using function $h(x_1, x_2) = u'(c_0 + x_1 + x_2)$. It is supermodular if the representative agent is prudent.

Proposition 2 *The presence of any positive first-order stochastic dependence in changes in consumption reduces the long-term risk-free rate if and only if the representative agent is prudent.*

This result confirms our intuition: positive FSD dependence in changes in consumption raises the riskiness of consumption at date T , without changing its expected value. Under prudence, this reduces the interest rate associated to maturity T . It tends to generate a downward-sloping yield curve.

⁵We can also prove that if h is supermodular, then condition (10) is satisfied if and only if $(\tilde{x}_1, \tilde{x}_2)$ exhibits positive FSC.

⁶See Joe (1997), Theorem 2.3.

It would have been more fashioned to define \tilde{x}_1 and \tilde{x}_2 as the (conditional) changes in *log* consumption over respectively subperiods $[0, t]$ and $[t, T]$. This change the nature of the comparative static exercise because the expectation of the *log* consumption is not the same as the *log* of the expectation of consumption. Because $\tilde{c}_T = c_0 e^{\tilde{x}_1 + \tilde{x}_2}$ when x_i denotes the change in *log* consumption, we can use Lemma 1 to obtain the following alternative result. Observe that $h(x_1, x_2) = u'(c_0 e^{x_1 + x_2})$ is supermodular if relative prudence $P(c) = -cu'''(c)/u''(c)$ is larger than unity.

Proposition 3 *The presence of any positive first-order stochastic dependence in changes in log consumption reduces the long-term risk-free rate if and only if relative prudence is larger than unity.*

Similarly, the long-term risk-free rate is increased by any negative FSD dependence if and only if relative prudence is larger than unity. Observe that when relative risk aversion is constant ($u(c) = c^{1-\gamma}/(1-\gamma)$), relative prudence is also constant and is equal to relative risk aversion plus one. Thus, CRRA implies that relative prudence is always larger than unity. When relative risk aversion is constant, positive (negative) FSD dependence in changes in *log* consumption always reduces (raises) the long-term risk-free rate relative to the benchmark of independent growth rates.

Corollary 1 *Suppose that u is a power function and that changes in log consumption $(\tilde{x}_1, \tilde{x}_2)$ have the same marginals and exhibit positive first-order stochastic dependence. It implies that the yield curve is decreasing: $r_T \leq r_t$.*

Proof: It is easy to check that the yield curve is flat ($r_t^i = r_T^i$) in the economy with the independent growth rate $(\tilde{x}_1, \tilde{x}_2^i)$. Because $r_t = r_t^i$ and $r_T \leq r_T^i$ from Proposition 3, we obtain that $r_T \leq r_t$ in the economy with the positively correlated growth rates $(\tilde{x}_1, \tilde{x}_2)$. ■

Notice that positive FSD dependence alone is not sufficient to obtain a decreasing yield curve. The above corollary relied on the additional assumption that relative risk aversion is constant. More generally, going back to Proposition 3, relative prudence must be larger than unity to obtain that positive FSD influences the long rate downwards. It is easy to exhibit utility functions that are concave but whose relative prudence is not larger than unity. For example, the simplest departure of CRRA with $u(c) = (c+k)^{1-g}/(1-g)$,

$k > 0$, implies a relative prudence $P(c) = (1+g)c/(c+k)$. For such a concave utility function, relative prudence tends to zero with c . At early stages of its development, this economy may have an upward sloping yield curve even if growth rates are positively FSD dependent. This comes from an implicit wealth increase. Observe that in spite of the fact that the dependence of $(\tilde{x}_1, \tilde{x}_2)$ does not affect the expected cumulative change in log consumption, the expected cumulative change in the *level* of consumption is increased by the presence of positive FSD dependence. This can be checked by using function $h(x_1, x_2) = e^{x_1+x_2}$ in the Lemma. This implicit increase in expected future incomes reduces the willingness to save for the long term, and it requires an increase in the corresponding interest rate. Therefore, one needs a sufficiently strong precautionary effect to dominate this opposite wealth effect. By Proposition 3, it requires that relative prudence be larger than unity.

4 Second-degree stochastic dependence

A natural extension of this work is to examine economies where the changes in consumption \tilde{x}_1 and \tilde{x}_2 are statistically related according to the positive second-degree stochastic dependence (SSD) property. This is the case when an increase in the first period change in consumption raises the risk associated to the second period change in consumption in the sense of Rothschild and Stiglitz (1970). In other words, the volatility of economic growth is increased after a boom, and it is reduced after a downturn. An example of such heteroskedastic process is $\tilde{x}_2 = \mu + \tilde{x}_1\tilde{\varepsilon}$, with $E\tilde{\varepsilon} = 0$ and $\tilde{\varepsilon}$ independent of \tilde{x}_1 .

Definition 2 *Consider a pair of random variables $(\tilde{x}_1, \tilde{x}_2)$. We say that there is a positive SSD dependence between \tilde{x}_1 and \tilde{x}_2 if $q(x_2 | x_1) = \int^{x_2} F_{2|1}(x_1, y)dy$ is non-decreasing in x_1 for all x_2 , and if $E[\tilde{x}_2 | x_1]$ is independent of x_1 .*

We want to determine the effect of such statistical relationship in changes in consumption over time on the long-term interest rate. As in the previous section, we compare an economy $(\tilde{x}_1, \tilde{x}_2)$ with positive SSD dependence with another one $(\tilde{x}_1, \tilde{x}_2^i)$ in which changes in consumption are serially independent with the same marginals. The following Lemma is helpful to solve this problem.

Lemma 2 Consider a twice differentiable bivariate function h . The following two conditions are equivalent:

1. For any pair of random variables $(\tilde{x}_1, \tilde{x}_2)$ that satisfies positive second-order dependence, we have that

$$Eh(\tilde{x}_1, \tilde{x}_2) \geq Eh(\tilde{x}_1, \tilde{x}_2^i) \quad (11)$$

2. $-\partial h / \partial x_2$ is supermodular, i.e., if $\partial^2 h / \partial x_2^2$ is non-increasing in x_1 .

Proof: See the appendix.

Notice that if we apply this lemma to function $h(x_1, x_2) = v(x_1 + x_2)$ for any function v with a convex first derivative, we obtain the result that, under positive SSD, $\tilde{x}_1 + \tilde{x}_2$ is a downside reduction in risk with respect to $\tilde{x}_1 + \tilde{x}_2^i$, as defined by Geiss, Menezes and Tressler (1980). It implies that these two sums have the same mean and the same variance, but the first has a larger skewness than the second. This is not a surprise since a downside reduction in risk is obtained by transferring zero-mean lotteries from low wealth states to larger wealth levels, as explained by Eeckhoudt, Gollier and Schneider (1995).

Applying this to the term structures given by (7) and (8) with $h(x_1, x_2) = u'(c_0 + x_1 + x_2)$, we obtain the following Proposition.

Proposition 4 *The presence of any positive second-degree stochastic dependence in changes in consumption raises the long-term risk-free rate if and only if the third derivative of the utility function is non-increasing.*

Observe that condition $u'''' \leq 0$ – which is sometimes referred to as ”temperance” – is quite natural. It is necessary for the intuitive property that absolute prudence $-u''' / u''$ is decreasing in wealth, as explained by Kimball (1990). Moreover, all CRRA functions satisfy this condition. There is a simple intuition for why a positive SSD in Δc should raise the equilibrium long-term rate. Indeed, a positive SSD implies an increase in skewness of $\tilde{c}_T = c_0 + \tilde{x}_1 + \tilde{x}_2$. When u'''' is negative, the increased skewness in \tilde{c}_T reduces $Eu'(\tilde{c}_T)$, which yields a reduction in the demand for the zero-coupon bond which matures at T . This raises its equilibrium rate of return. The assumption that the fourth derivative of the utility function is negative is compatible

with decreasing prudence (Kimball (1990)), and with risk vulnerability (Gollier and Pratt (1996)), two very intuitive behavioral assumptions.

As in the previous section, we could have defined \tilde{x}_1 and \tilde{x}_2 as the changes in log consumption. We should then use Lemma 2 with $h(x_1, x_2) = u'(c_0 e^{x_1+x_2})$. This yields the following result.

Proposition 5 *Suppose that u is four times differentiable. The presence of any positive second-order stochastic dependence in changes in log consumption raises the long-term risk-free rate if and only if $f(c) = u''(c) + 3cu'''(c) + c^2u''''(c)$ is uniformly negative.*

When relative risk aversion is constant, $f(c)$ equals $-\gamma^3 c^{-\gamma-1}$ which is uniformly negative. This implies that a positive SSD dependence in $\Delta \ln c$ always raises the long-term interest rate for that family of utility functions. This tends to generate an upward-sloping yield curve. The proof of the following corollary parallels the one of Corollary 1 and is therefore skipped.

Corollary 2 *Suppose that u is a power function and that changes in log consumption $(\tilde{x}_1, \tilde{x}_2)$ have the same marginals and exhibit positive second-order stochastic dependence. It implies that the yield curve is increasing: $r_T \geq r_t$.*

Notice that condition $f \geq 0$ in Proposition 5 adds the two terms u'' and $3cu'''$ to condition $u'''' < 0$ in Proposition 4. The first additional term is due to the fact that making changes in log consumption serially independent raises the expected consumption at T . This reinforces the initial reason for a longer long-term rate. Also, it yields an increase in the second moment of \tilde{c}_T . Under prudence, this tends to reduce the long-term rate. This explains the opposite term $3cu''' > 0$ in the definition of function f . As said above, the two negative terms must always dominate this positive term in the case of CRRA.

5 Relations with the existing literature on the term structure

Our aim in this section is not to provide a survey of the enormous existing literature on the term structure of interest rates. Rather, we want here to

illustrate our results by comparing them to those of the two most famous time series models used in this literature: Vasicek (1977) and Cox, Ingersoll and Ross (1985a,b). In most existing models of the term structure, the state variable is the instantaneous interest rate r_0 . For example, in the model of Vasicek (1977), the time series model for the stochastic discount factor $\Lambda(t) = e^{-\delta t} u'(c_t)/u'(c_0)$ takes the following form:

$$\begin{aligned}\frac{d\Lambda}{\Lambda} &= -r_0 dt - \sigma_\Lambda dz \\ dr_0 &= \phi(\bar{r} - r_0)dt + \sigma_r dz.\end{aligned}$$

The term structure is then obtained by rewriting the equilibrium condition (1) as $r_t = -t^{-1} \ln E\Lambda_t$. Parameter σ_r is the conditional volatility of the instantaneous interest rate. Parameter ϕ controls mean reversion: if $\phi = 0$, the instantaneous risk-free rate r_0 exhibits no tendency to return to any specific value. When $\phi > 0$, the instantaneous rate r_0 is expected to return to its mean \bar{r} at rate ϕ . With a typical value of $\phi = 0.3 \text{ year}^{-1}$,⁷ this yields a half-life time of 2.3 years for a shock on the instantaneous interest rate. Assuming that $u'(c) = c^{-\gamma}$, this model can be rewritten as⁸

$$\begin{aligned}d \ln c &= \frac{r_0 - \delta + 0.5\sigma_\Lambda^2}{\gamma} dt + \frac{\sigma_\Lambda}{\gamma} dz \\ dr_0 &= \phi(\bar{r} - r_0)dt + \sigma_r dz.\end{aligned}\tag{12}$$

Campbell (1986) examines a more general (discrete) version of this model in which the first difference of the log endowment follows a univariate, stationary stochastic process with a constant drift. We recognize in equation (12) various elements affecting the yield curve. First, we observe that the conditional volatility of the growth rate of consumption is a constant σ_Λ/γ . This excludes the existence of SSD dependence. Second, when $\phi \neq 0$, there is a deterministic component in the expectations about the future growth of

⁷For example, Backus, Foresi and Telmer (1998) consider $\phi = 0.024 \text{ month}^{-1}$.

⁸By Ito's Lemma, the reader can check that

$$-\delta dt - \gamma d \ln c = d \ln \Lambda = \frac{d\Lambda}{\Lambda} - 0.5\sigma_\Lambda^2 dt.$$

the economy. When the current level of r_0 is below \bar{r} , one anticipates an accelerating economic growth, which makes the yield curve increasing for short and medium maturities. In fact, when there is no serial dependence ($\sigma_r = 0$), using equation (6) yields

$$r_t = \bar{r} + (r_0 - \bar{r}) \frac{1 - e^{-\phi t}}{\phi t} \quad (13)$$

The mean yield curve, which is obtained by taking $r_0 = \bar{r}$, is completely flat in that case. Thus, a non-zero slope to the mean yield curve can be obtained only by introducing some permanency to shocks on $\ln c$. Indeed, when $\sigma_r \neq 0$ and has the same sign as σ_Λ , there is positive FSD dependence in the time series of $\Delta \ln c$. As claimed by Proposition 3, this tends to reduce the long-term interest rates, thereby yielding a reduction in the slope of the yield curve. This is confirmed by the analytical solution obtained by Vasicek (1977) which adds a third term v_t in the right-hand side of (13), with v_t being negative and decreasing in t when $\sigma_r \sigma_\Lambda$ is positive.⁹ In order to explain the upward-sloping mean yield curve, as documented for example by Backus, Foresi and Telmer (1998), one needs to have negative FSD dependences in $\Delta \ln c$. Assuming without loss of generality that σ_r is positive, this requires that σ_Λ/γ be sufficiently negative. This sheds light on why "a little experimentation tells us that σ_Λ governs the average slope of the yield curve, with negative values required to produce an upward slope we observe in the data" (Backus, Foresi and Telmer (1998)).

The Cox-Ingersol-Ross (CIR) model adds a square root terms in the volatility, which makes it time-varying:

$$\begin{aligned} d \ln c &= \frac{r - \delta + 0.5\sigma_\Lambda^2}{\gamma} dt + \frac{\sigma_\Lambda \sqrt{r}}{\gamma} dz \\ dr &= \phi(\bar{r} - r) dt + \sigma_r \sqrt{r} dz. \end{aligned} \quad (14)$$

This implies that a positive shock on the state variable r increases both the short term expected growth of the log consumption and its future volatility.

⁹Vasicek (1977) obtained

$$v_t = \left[\frac{\sigma_r^2}{2\phi^2} + \frac{\sigma_r \sigma_\Lambda}{\phi} \right] \left[\frac{1 - e^{-\phi t}}{\phi t} - 1 \right] + \frac{\sigma_r^2}{4\phi^3} \frac{(1 - e^{-\phi t})^2}{t}.$$

This yields a positive SSD dependence of $\Delta \ln c$. This captures the fact that higher interest rates seem to be more volatile. As claimed in Proposition 5, this positive SSD in $\Delta \ln c$ generates an unambiguous increase in the long-term rate compared to what is obtained in Vasicek's model. However, as explained for example by Backus, Foresi and Telmer (1998), the sophistication introduced in the CIR model does not change the shape of the yield curve markedly.

We hereafter focus on quantifying the effect of FSD dependences. What do we know about the time series of changes in log consumption? Following Cochrane (1988) and Cogley (1990), let us define the variance ratio as

$$V(t) = \frac{t^{-1} \text{Var}(\ln \tilde{c}_t - \ln c_0)}{\text{Var}(\ln \tilde{c}_1 - \ln c_0)}.$$

The variance ratio associated to time horizon t equals the variance of the change in log consumption at horizon t divided by t times the variance of one year changes in log consumption. $V(t)$ provides a relative measure of the uncertainty associated to $\ln \tilde{c}_t | c_0$. In the case of serially independent $\Delta \ln c$, this variance ratio is uniformly equal to unity. $V > 1$ indicates a positive FSD dependence in $\Delta \ln c$, whereas $V < 1$ indicates a negative FSD dependence. Cochrane (1988) estimated $V(t)$ for $t = 1, \dots, 30$ by using data on the log real per capita GNP in the United States, 1869-1986. Figure 1 summarizes his estimates. The per period risk attached to time horizons less than 3 years is increasing. This comes from the positive serial dependence of growth at high frequency. On the contrary, V is decreasing in t for time horizons longer than 3 years. It tends to roughly one-third. Long horizons entail only one-third per period risk than short horizons, when risk is measured by the variance of log consumption. This means that shocks to U.S. GNP are mostly temporary. Thus, in the U.S. context, $\Delta \ln c$ exhibits negative FSD dependences for long maturities. According to Proposition 3, assuming CRRA, this should generate an upward-sloped unconditional yield curve.

Taking equation (5) as an approximation because $\tilde{c}_T | c_0$ is generally not lognormal, we can use Cochrane's V ratio to obtain the following formula for the yield curve:

$$r_t = \delta + \gamma\mu - 0.5\gamma^2\sigma^2V(t), \tag{15}$$

where $\sigma = \sqrt{\text{Var} \ln \tilde{c}_1 | c_0}$ is estimated to be 6.1% per year by Cochrane. Let us fix the expected growth rate of the economy to $\mu = 1.8\%$ per year, which is

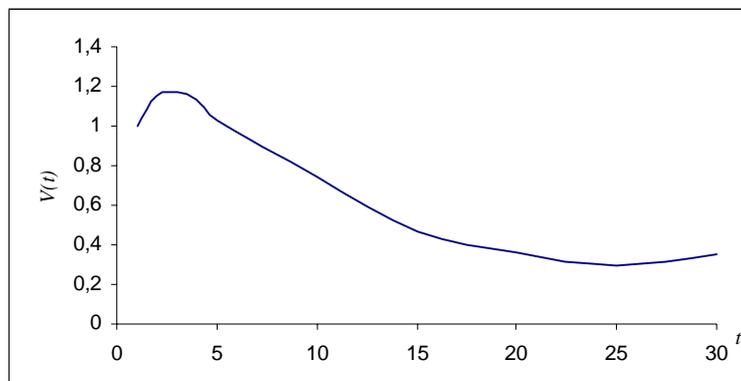


Figure 1: The variance ratio for the log real per capita GNP, 1869-1986. (Source: Cochrane (1988)).

the average growth rate of real per capita consumption in the United States over the period 1889-1978 (Kocherlakota (1996)). In Figure 2, we draw the yield curve $r_t - \delta$ computed from equation (15) for four different degrees of relative risk aversion: $\gamma = 1, 2, 4$ and 6. The upward-sloping shape of the average yield curve is familiar. Using U.S. monthly data from January 1952 to February 1991, Backus, Foresi and Telmer (1998) estimated the mean 1-month yield to be 5.314%, going up to 6.693% for the yield corresponding to a 10-year maturity.

Cogley (1990) showed that the pattern of the variance ratio exhibits much differences across countries. In fact, the evidence indicates that the relative stability of long-term growth is unique to the United States. Using annual real per capita GDP, 1871-1985, he computed the variance ratio $V(20)$ for a twenty years horizon. He found 0.77 for Canada, which means that, as in the U.S. but at a smaller degree, this country should have a mean 20-year maturity yield that is larger than the short-term yield. He also found 0.97 for Sweden, 1.03 for the United Kingdom, and 1.09 for Denmark. The yield curve should be almost flat in these countries. But he also obtained 1.4 for Australia, 1.84 for France and 2.02 for Italy. In these countries, the per-period growth risk is increasing with time horizon. It implies that the long-term interest rate should be smaller than the short-term one. For

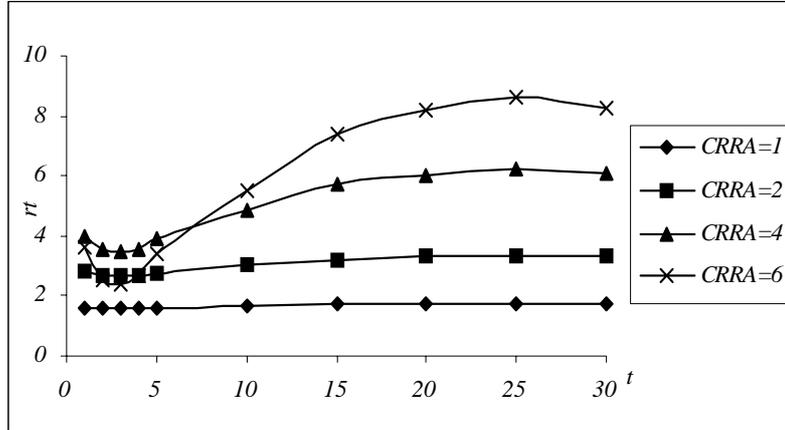


Figure 2: The average yield curve using Cochrane’s ratio.

France, using Maddison (1991), we estimated $\mu = 1.97\%$ and $\sigma = 8.05\%$. For $\gamma = 2$, it makes a risk-free rate r_t equaling $\delta + 2.66\%$ for the short term, and $\delta + 1.58\%$ for the long run. For $\gamma = 4$, it generates $\delta + 2.76\%$ and $\delta - 1.54\%$ respectively for the short term and for the long term.

6 What about far-distant maturities?

As explained in the introduction, our aim is to determine whether the discount rates that we should use for very long maturities (in the hundreds and thousands years) are smaller than the discount rates used for the more standard maturities considered in finance and that can be observed on financial markets. Whereas most of the existing literature provides a positive analysis of the term structure, our aim in this section is more normative. As explained earlier in this paper, the evidence that there is some mean-reversion in consumption growth with a half-life of 2.3 years is important to explain the shape of the yield curve observed on financial markets. It is of course irrelevant to determine the rate that we should use to discount the cash-flows associated with the consequences of, say, global warming or nuclear wastes in 200 years. In this section, we show how the results presented in sections

3 and 4 are useful to make recommendations for such long time horizons.

We examine two possible dynamic processes governing the long-term growth of the economy. The first one involves Poisson jumps, whereas the other one exhibits some parameter uncertainty.

6.1 Two-state jumps in the growth of consumption

We reproduced in Table 1 the estimation by DeLong (2004) of the worldwide GDP per capita over the last 7 millennia. A striking feature of the economic growth that we observed over this essential period of homo sapiens is the jump in the average growth rate that has been experienced at the end of the eighteenth century. The debate on what Society should do for the future is much influenced by this fact, and by the idea that in the future, economic growth could go back to its secular zero level. The arguments for and against the existence of such risk are based on the way we interpret what happened at the end of the eighteenth century. Models abound that rely on human capital, availability of natural resources, specialization of tasks,... We will not go into this controversy. As earlier in this paper, the economic growth process will not be endogenized.

Year	GDP per capita	Annualized growth rate
-5000	130	
-1000	160	0.005 %
1	135	-0.017 %
1000	165	0.020 %
1500	175	0.012 %
1800	250	0.119 %
1900	850	1.224 %
1950	2030	1.741 %
1975	4640	3.307 %
2000	8175	2.265 %

Table 1: Worldwide GDP per capita (in year-2000 international dollars) and annualized growth rate (in %). Source: DeLong (2004)

We assume that the economy can be in one of two observable states. In the low-growth state, the drift of log consumption is μ_1 and its volatility is σ . In the high-growth state, the drift of log consumption is $\mu_2 > \mu_1$, but

the volatility remains the same. The economy switches from one state to the other each time a Poisson event occurs. In discrete time, the model is written as

$$\begin{aligned}\ln \tilde{c}_{t+1} &= \ln c_t + \mu_t + \sigma \tilde{\varepsilon}_t \\ \mu_{t+1} &= (\mu_t, 1 - \pi; \mu'_t, \pi),\end{aligned}$$

where μ'_t is μ_2 if $\mu_t = \mu_1$, otherwise $\mu'_t = \mu_1$. We assume that $\tilde{\varepsilon}_t$ is standard normal with no serial dependency. When the drift is μ_1 (resp. μ_2) in period t , there is a probability π that it will switch to drift μ_2 (resp. μ_1) in the next period. We suppose that the probability of switching is very small, which implies that there is a lot of persistence in shocks. Of course, this yields FSD dependence in growth rates. There is positive FSD dependence as soon as π is less than 0.5. Thus, applying Proposition 3 implies that the far-distant discount rate is smaller than in the equivalent economy with a serially independent growth process. In Figure 3, we describe the yield curve in the following economy: In the high-growth state, the drift of log consumption is $\mu_2 = 2\%$ per year, whereas it is $\mu_1 = 0\%$ in the low-growth state. The critical parameter is the switching probability π per year. We assume that the switching probability π per year is 1%. Relative risk aversion is assumed to be constant and equal to $\gamma = 2$. We consider a rate δ of pure preference for the present and a conditional volatility σ of log consumption such that $\delta - 0.5\gamma^2\sigma^2 = 0$.¹⁰ We assume that the economy is currently in the high-growth state. Figure 3 also describes the yield curve in the equivalent economy with a serially independent growth process. The details of the computational method generating these two yield curves are given in the appendix. Whereas the discount rate is 4% in the "short" term, it goes down to 1.03% for cash-flows with a 500-year maturity.

6.2 A model with parameter uncertainty

As invoked in the so-called Peso-problem, the absence of sufficiently large data to estimate the long-term growth process of the economy implies that the parameters controlling the growth process are uncertain and subject to learning in the future. Weitzman (2004) shows that, under CRRA and geometric Brownian motion, the uncertainty surrounding the true drift of log

¹⁰As usual, considering other values for δ and σ would just shift the yield curve vertically.

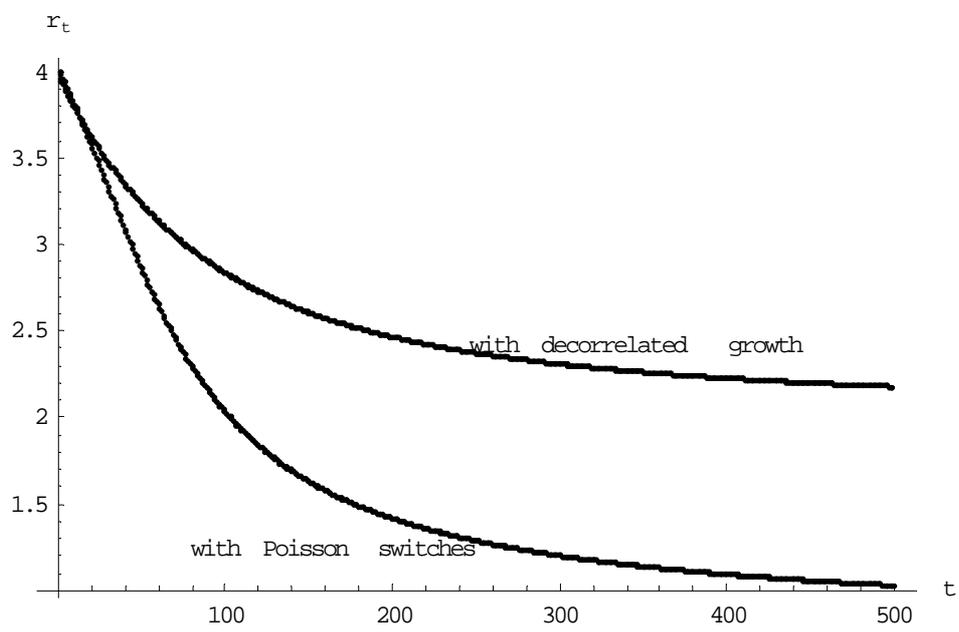


Figure 3: The yield curve with a two-state Poisson switch process for the growth of log consumption, in the high-growth state.

consumption justifies selecting a smaller rate to discount distant cash-flows. In this section, we explain this phenomenon and we provide a more general model than in Weitzman (2004). The intuition for why the uncertainty surrounding the drift of the growth process justifies selecting a smaller long discount rate is immediate from Proposition 2. Indeed, the observation of a high growth in the short run induces the representative agent to revise her expectations about the distribution of growth upwards. Thus, Bayesian learning generates positive FSD dependence in the perceived growth process. This magnifies the long-term risk, thereby inducing the prudent representative agent to make more effort for the distant future. As shown by Proposition 2, this result requires no other restriction on preferences than prudence.

We suppose that the growth process is stationary. Let $\tilde{x}(\theta)$ denote the per-period change in log consumption, conditional to parameter θ . The current prior beliefs of the representative agent are described by the distribution of random variable $\tilde{\theta}$. Under CRRA preferences, the current yield curve takes the following form:¹¹

$$r_t = \delta - \frac{1}{t} \ln E\alpha(\tilde{\theta})^t, \quad (16)$$

where function α is defined as

$$\alpha(\theta) = Ee^{-\gamma\tilde{x}(\theta)}. \quad (17)$$

Using Jensen's inequality, we directly get the following result, which is related to Proposition 3.

Proposition 6 *Suppose that the representative agent has CRRA preferences, and that the process of log consumption is stationary with an unknown parameter θ . Under such circumstances, the socially efficient discount rate r_t is non-increasing with time horizon t . It tends to the smallest possible rate $\min_{\theta} [\delta - \ln \alpha(\theta)]$ when t tends to infinity.*

Proof: Observe first that function $g(x) = x \ln x$ is convex. Then, using Jensen's inequality, we have that

$$t^2 E \left[\alpha(\tilde{\theta})^t \right] \frac{\partial r_t}{\partial t} = \left[E\alpha(\tilde{\theta})^t \right] \left[\ln E\alpha(\tilde{\theta})^t \right] - E \left[\alpha(\tilde{\theta})^t \ln \alpha(\tilde{\theta})^t \right]$$

¹¹In the economy with serial independence, we just have $\hat{r}_t = \delta - \ln E\alpha(\tilde{\theta})$.

is nonpositive. Thus, r_t is non-increasing in t . Moreover, as is well-known, when t tends to infinity, $\left[E\alpha(\tilde{\theta})^t\right]^{1/t}$ tends to $\max_{\theta} \alpha(\theta)$, which implies that r_t tends to $\delta - \ln \max_{\theta} \alpha(\theta)$. ■

Notice that r_t is strictly decreasing in t as soon as there exists two values of the parameter, θ and θ' such that $\alpha(\theta) \neq \alpha(\theta')$. This result and its proof is reminiscent – but is conceptually different – of a recommendation in Weitzman (1998) for why “the far-distant future should be discounted at its lowest possible rate”. Notice also that the above result does not require any condition on the distribution of changes in log consumption $\tilde{x}(\theta)$, or on the prior distribution of parameter $\tilde{\theta}$. Weitzman (2004) assumes that $\tilde{x}(\theta)$ is normal with a known volatility σ , which implies that $\alpha(\theta) = \exp(-\gamma\mu(\theta) + 0.5\gamma^2\sigma^2)$. It implies in turn that the discount rate tends to $\delta - 0.5\gamma^2\sigma^2 + \gamma \min_{\theta} \mu(\theta)$. Because he also assumes that $\mu(\tilde{\theta})$ is normally distributed, the discount rate goes to minus infinity for large maturities under this specification!¹²

More realistic specifications of the per-period growth process and/or of the prior beliefs are thus welcomed. Equations (16) and (17) provide this simple and flexible framework. Consider for example the following numerical illustration. The relative risk aversion of the representative agent equals $\gamma = 2$. The change in log consumption is normal with conditional standard deviation $\sigma(\theta)=6.1\%$, whereas we assume that $\delta - 0.5\gamma^2\sigma^2 = 0$. The drift μ is unknown, but it is either 3% or 0%. The prior belief is that there is a 2/3 probability that the true drift is 3%, yielding an expected drift of 2%. If it would be 3% for sure, the yield curve would be flat at 6%, whereas it would be equal to 0% in the low-growth scenario. In Figure 4, we draw the yield curve given the current parameter uncertainty. The learning process induces Society to use today a 0.22% rate per year to discount cash-flows realized in 500 years, whereas a discount rate of 4.0% per year is used for immediate benefits and costs.

7 Conclusion

A correct assessment of how much Society should invest for its own future is central to economic analysis. Many of us are now cooperating with various

¹²If $\mu(\tilde{\theta})$ is normally distributed with mean μ and variance σ_0^2 , we obtain that $r_t = \delta + \gamma\mu - 0.5\gamma^2(\sigma^2 + t\sigma_0^2)$ which decreases linearly with the time horizon.

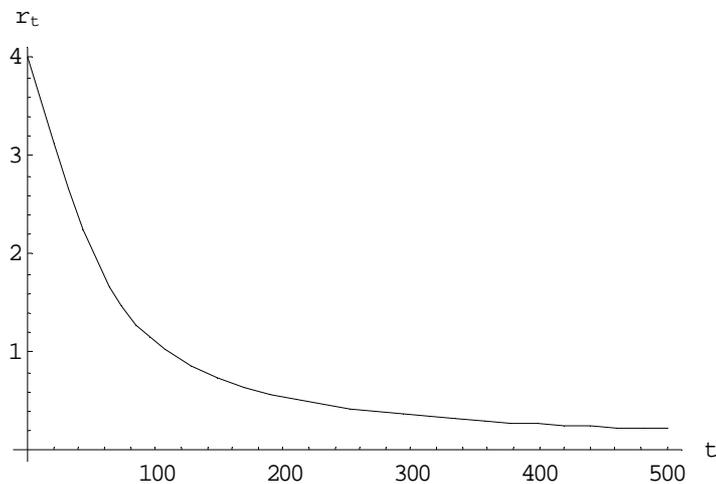


Figure 4: The yield curve with parameter uncertainty.

organizations to analyze environmental projects whose costs and benefits are spread over hundreds of years, in particular those linked to global warming and nuclear waste disposals. We know that the most important parameter when using cost-benefit analysis for such long-lasting projects is by far the discount rate. We as a profession have not been very good in proposing an agreed-upon discount rate for the long term. Weitzman (2001) asked to more than 2000 professional Ph.D.-level economists about their own recommendation for the discount rate to be used for far distant real cash-flows. He reported a large disagreement on this matter. Moreover, he obtained a sample mean at around 4% per year, which is quite larger than the secular post-industrial-revolution real short-term interest rate of 1% (Kocherlakota (1996)). Economists seem to favor an upward-sloping discount yield curve.

The main message of the paper is that the shape of the term structure of discount rates depends essentially on our view about how the uncertainty on future aggregate consumption evolves with the time horizon. If this uncertainty increases at a rate larger than what would be obtained by a pure random walk for the per-period growth rate, the notion of prudence justifies using a downward-sloping term structure. This is the case when per-period growth rates exhibit positive first-degree stochastic dependence, as is the

case with persistent shocks to growth rates, or when the drift of aggregate consumption is unknown. Our calibrations induce us to recommend using an average yearly discount rate of 4% for short-term cash-flows, and a yearly discount rate between 1% and 2% for time horizons exceeding 400 years.

REFERENCES

- Arrow, K.J., W.R. Cline, K.-G. Maler, M. Munasinghe, R. Squitieri and J.E. Stiglitz, (1996), Intertemporal equity, discounting and economic efficiency, in *Climate Change 1995 - Economic and Social Dimensions of Climate Change*, eds J.P. Bruce, H. Lee and E.F. Haites, Cambridge University Press, Cambridge U.K..
- Backus, D., S. Foresi and C. Telmer, (1998), Discrete-time models of bond pricing, NBER Working Paper 6736.
- Barsky, R.B., (1989), Why don't the prices of stocks and bonds move together?, *American Economic Review*, 79, 1132-1145.
- Campbell, (1986), Bond and stock returns in a simple exchange model, *Quarterly Journal of Economics*, 101, 785-804.
- Cochrane, J.H., (1988), How big is the random walk in GNP?, *Journal of Political Economy*, 96, 893-920.
- Cochrane, J., (2001), *Asset Pricing*, Princeton University Press.
- Cogley, T., (1990), International evidence on the size of the random walk in output, *Journal of Political Economy*, 98, 501-518.
- Cox, J., Ingersoll, J., and S. Ross, (1985a), A theory of the term structure of interest rates, *Econometrica*, 53, 385-403.
- Cox, J., Ingersoll, J., and S. Ross, (1985b), An intertemporal general equilibrium model of asset prices, *Econometrica*, 53, 363-384.
- Breeden, D.T., (1986), Consumption, production, inflation, and interest rates: A synthesis, *Journal of Financial Economics*, 16, 3-40.
- DeLong, B.J., (2004), Chapter 5: The reality of economic growth: History and prospect, <http://www.j-bradford-delong.net>.
- Eeckhoudt, L., C. Gollier. and T. Schneider, (1995), Risk aversion, prudence and temperance: A unified approach, *Economics Letters*, 48, 331-336.

- Estrella, A., and G.A. Hardouvelis, (1991), The term structure as a predictor of real economic activity, *Journal of Finance*, 46, 555-576.
- Geiss, C. , C. Menezes and J. Tressler, (1980), Increasing Downside Risk, *American Economic Review*, 70, 5, 921-931.
- Gollier, C., (2002a), Discounting an uncertain future, *Journal of Public Economics*, 85, 149-166.
- Gollier, C., (2002b), Time horizon and the discount rate, *Journal of Economic Theory*, 107, 463-473.
- Gollier, C. and J.W. Pratt, (1996), Risk vulnerability and the tempering effect of background risk, *Econometrica*, 64, 1109-1124.
- Groom, B., P. Koundouri, E. Panopoulou and T. Pantelidis, (2004), Model selection for estimating certainty equivalent discount, mimeo, UCL, London.
- Hansen, L. and K. Singleton, (1983), Stochastic consumption, risk aversion and the temporal behavior of assets returns, *Journal of Political Economy*, 91, 249-265.
- Joe, H., (1997), *Multivariate models and dependence concepts*, Chapman and Hall/CRC.
- Kimball, M.S., (1990), Precautionary savings in the small and in the large, *Econometrica*, 58, 53-73.
- Kocherlakota, N.R., (1996), The Equity Premium: It's Still a Puzzle, *Journal of Economic Literature*, 34, 42-71.
- Lehmann, E.L., (1966), Some concepts of dependence, *Annals of Mathematical Statistics*, 37, 1137-1153.
- Leland, H., (1968), Savings and uncertainty: The precautionary demand for savings, *Quarterly Journal of Economics*, 45, 621-36.
- Lucas, R., (1978), Asset prices in an exchange economy, *Econometrica*, 46, 1429-46.
- Maddison, A., (1991), *Phases of Economic Development*, Oxford Economic Press.

- Mankiw, G., (1981), The permanent income hypothesis and the real interest rate, *Economic Letters*, 7, 307-311.
- Milgrom, P., (1981), Good news and bad news: Representation theorems and applications, *Bell Journal of Economics*, 12, 380-91.
- Newell, R., and W. Pizer, (2003), Discounting the benefits of climate change mitigation: How much uncertain rates increase valuations?, *Journal of Environmental Economics and Management*, 46 (1), 52-71.
- Piazzesi, M., (2005), Affine term structure models, in *Handbook of Financial Econometrics*, Y. Ait-Sahalia and L.P. Hansen eds, Elsevier.
- Portney, P.R., and J. P. Weynant, eds., (1999), *Discounting and intergenerational equity*, Resources for the future, Washington, D.C..
- Rothschild, M. and J. Stiglitz, (1970), Increasing risk: I. A definition, *Journal of Economic Theory*, 2, 225-243.
- Shaked, M., and J.G. Shanthikumar, (2007), *Stochastic Orders*, Springer series in statistics.
- Tchen, A.H., (1980), Inequalities for distributions with given marginals, *Annals of Probability*, 8, 814-827.
- Vasicek, O., (1977), An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5, 177-188.
- Weitzman, M.L., (1998), Why the far-distant future should be discounted at its lowest possible rate?, *Journal of Environmental Economics and Management*, 36, 201-208.
- Weitzman, M.L., (2001), Gamma discounting, *American Economic Review*, 91, 260-271.
- Weitzman, M.L., (2004), Statistical discounting of an uncertain distant future, mimeo, Harvard University.

Proof of Lemma 1

Define function K as: $K(x_1, x_2) = F_{2|1}(x_1, x_2) - F_2(x_2)$, where $F_{2|1}$ is the conditional distribution of \tilde{x}_2 and F_2 is its marginal distribution. For $2 \implies 1$, we need to prove that

$$X = Eh(\tilde{x}_1, \tilde{x}_2) - Eh(\tilde{x}_1, \tilde{x}_2^i) = \iint h(x_1, x_2) d[F(x_1, x_2) - F_1(x_1)F_2(x_2)]$$

is positive. For any given x_1 , integration by parts yields

$$\int h(x_1, x_2) d[F(x_1, x_2) - F_1(x_1)F_2(x_2)] = - \int \frac{\partial h(x_1, x_2)}{\partial x_2} K(x_1, x_2) dx_2 dF_1(x_1). \quad (18)$$

It implies that

$$X = \int \left[\int -\frac{\partial h(x_1, x_2)}{\partial x_2} K(x_1, x_2) dF_1(x_1) \right] dx_2,$$

or equivalently,

$$X = \int E \left[-\frac{\partial h(\tilde{x}_1, x_2)}{\partial x_2} K(\tilde{x}_1, x_2) \right] dx_2. \quad (19)$$

Observe now that for any x_2 , $-\partial h/\partial x_2$ is decreasing in x_1 because h is supermodular. Moreover, K is decreasing in x_1 for all x_2 by definition of positive FSD. Therefore for any x_2 , the covariance rule¹³ implies that

$$E \left[-\frac{\partial h(\tilde{x}_1, x_2)}{\partial x_2} K(\tilde{x}_1, x_2) \right] \geq E \left[-\frac{\partial h(\tilde{x}_1, x_2)}{\partial x_2} \right] E [K(\tilde{x}_1, x_2)] = 0.$$

Since the integrand in (19) is positive for all x_2 , so is the integral X . This proves that $2 \Rightarrow 1$.

For $1 \implies 2$, suppose by contradiction that $-\partial h/\partial x_2$ be increasing in x_1 in a neighborhood A of some (\bar{x}_1, \bar{x}_2) . Using a pair of random variables satisfying positive FSD whose support is in A would generate $X \leq 0$, a contradiction. ■

¹³ $Ef(\tilde{x})g(\tilde{x}) \geq Ef(\tilde{x})Eg(\tilde{x})$ for all \tilde{x} if $f'(x)g'(x)$ is nonnegative for all x . See for example Gollier (2001, section 6.4) for a formal proof.

Proof of Lemma 2

We limit the proof to sufficiency. Let $k(x_1, x_2)$ denote $\int^{x_2} K(x_1, y)dy$. Integrating by parts the integral in the right-hand side of equation (18) yields

$$\int h(x_1, x_2)dK(x_1, x_2) = -\frac{\partial h(x_1, x_2)}{\partial x_2} \lim_{y \rightarrow \infty} k(x_1, y) + \int \frac{\partial^2 h(x_1, x_2)}{\partial x_2^2} k(x_1, x_2) dx_2 \quad (20)$$

for all x_1 . By construction, we have that

$$\lim_{y \rightarrow \infty} k(x_1, y) = \int (F_{2|1}(x_1, y) - F_2(y)) dy = E[\tilde{x}_2 | x_1] - E[\tilde{x}_2] = 0$$

since the expectation of \tilde{x}_2 is assumed to be independent of x_1 . Thus we can use (20) to write

$$X = Eh(\tilde{x}_1, \tilde{x}_2) - Eh(\tilde{x}_1, \tilde{x}_2^i) = \int E \left[\frac{\partial^2 h(\tilde{x}_1, x_2)}{\partial x_2^2} k(\tilde{x}_1, x_2) \right] dx_2.$$

Positive SSD means that k is increasing in x_1 for all x_2 . Because $\partial^2 h / \partial x_2^2$ is decreasing in x_1 by assumption, the covariance rule applied for each possible x_2 implies that

$$X \leq \int E \left[\frac{\partial^2 h(\tilde{x}_1, x_2)}{\partial x_2^2} \right] E[k(\tilde{x}_1, x_2)] dx_2.$$

Because $E k(\tilde{x}_1, x_2)$ is zero for all x_2 by construction, we obtain that X is nonpositive. ■

Computational method used to draw Figure 3

We use the iterative method presented for example by Backus, Foresi and Telmer (1998) to compute the term structure in the two-state economy with Poisson switches. Let b_τ^t denote the price at date τ of a zero-coupon bond with maturity t . The CRRA representative agent should be indifferent to increase her demand for such bond, yielding the following equilibrium condition:

$$b_\tau^{t+1} = E e^{-\delta} \left(\frac{\tilde{c}_{\tau+1}}{c_\tau} \right)^{-\gamma} b_{\tau+1}^t. \quad (21)$$

Applying this condition for $t = 1$ and observing that $b_{t+1}^0 = 1$ yields

$$r_1 = -\ln b^1 = \delta - \frac{1}{2}\gamma^2\sigma^2 + \gamma\mu,$$

with $b_\tau^t = b^t(\mu_\tau)$. Iterating (21) forward generates the following formula:

$$\begin{aligned} f^1(\mu) &= 0 \\ f^{t+1}(\mu) &= -\frac{1}{\gamma} \ln \left[(1 - \pi) e^{-\gamma(\mu + f^t(\mu))} + \pi e^{-\gamma(\mu' + f^t(\mu'))} \right] \\ \ln b^t &= -t(\delta - 0.5\gamma^2\sigma^2) - \gamma(\mu + f^t(\mu)) \end{aligned}$$

and, finally,

$$r_n = -\frac{1}{t} \ln b^t = \delta - \frac{1}{2}\gamma^2\sigma^2 + \gamma \frac{\mu + f^t(\mu)}{t}.$$

The analysis of the economy with no serial dependency is organized as follows. Let \tilde{y}_t be distributed as the change in log consumption in period t conditional to a high-growth state at date 0. Thus, conditional to $\mu(t = 0) = \mu_2$, $(\tilde{y}_1, \tilde{y}_2, \dots)$ is an independent process. It happens that \tilde{y}_t is distributed as $(\mu_2, 1 - \pi_t; \mu_1, \pi_t) + \sigma\tilde{\varepsilon}$ where π_t is the probability of a odd number of Poisson events between 0 and t . The pricing formula in the high-growth state is then

$$r_t = \delta - \frac{1}{t} \ln E \prod_{i=1}^t \tilde{y}_i^{-\gamma} = \delta - \frac{1}{2}\gamma^2\sigma^2 - \frac{1}{t} \sum_{i=1}^t \ln [(1 - \pi_i) e^{-\gamma\mu_2} + \pi_i e^{-\gamma\mu_1}].$$

An equivalent pricing formula is easily obtained for the low-growth state.